



ISSN No: 2319-5886

International Journal of Medical Research & Health Sciences, 2016, 5, 9S:321-325

Taxonomy level of pediatrics exams and psychometric characteristics of MCQ test in year 2014

¹Vakili Zarichehr, ²Sharif Mohammadreza, ³Sayyah Mansour and ⁴Khosravi Gholamreza

¹Department of Pathology, Affiliated Faculty Member of Education Development Center, Kashan University of Medical Sciences, Kashan, I.R. Iran

²Department of Pediatrics, Affiliated Faculty Member of Education Development Center, Kashan University of Medical Sciences, Kashan, I.R. Iran

³Affiliated Faculty Member of Education Development Center, Kashan University of Medical Sciences, Kashan, I.R. Iran

⁴Orthopedic Department, Affiliated Faculty Member of Trauma Research Center, Kashan University of Medical Sciences, Kashan, I.R. Iran

Corresponding Author: Gkhosravi@yahoo.com

ABSTRACT

Educational assessment not only needs to examine the test items making up the test but also to evaluate the level of learning of test takers. The purpose of this research was to examine the multiple choice items indices and to determine the taxonomy of pediatric exams of a medical university. This cross sectional descriptive research included 150 MCQ items making up the exams of pediatric group test for residents in Kashan University of Medical Sciences and taxonomy level of the test was determined. Computer software was used to determine the psychometric indices of MCQ test and expert opinion was used to determine the taxonomy of the test. The result of analysis indicated that 42.7 percent of the question was measuring recall, 31.3 percent were at the comprehension level, and 26 percent were targeting application. Chi square test result indicated that there was a significant difference between the proportion of the levels ($P=0.03$). In addition, a significant difference was also found between the proportion of very difficult and easy test items ($P=0.024$). Based on the result of this research, it seems like MCQ tests used to exam the resident in the pediatric group need further evaluations and more careful examination.

Keywords: Taxonomy, item analysis, pediatric assistance

INTRODUCTION

Academic assessment is an integral part of all educational institutions. Nearly all academic institutions employ different forms of written or verbal exams to evaluate their students. A good assessment is necessary for different purposes and has significant impact on students' learning; whereas a poor or faulty assessment has many negative consequences such as unfair grades, loss of motivation and students objections to name a few(1, 2). For this reason, in situation when test such as MCQs are used, item analysis is performed and psychometric characteristic of the test is determined (3-7). In these types of test descriptions, item discrimination, item difficulty, internal consistency, and distracters to the response alternatives are examined. There is certain criterion to judge the adequacy of the test items to be included within the collection of test items. The availability of computer software to perform the item analysis facilitates the procedure. However, it seems like having good test items alone are not sufficient to conclude that the test is measuring all aspects of learning objectives. In spite of the significance of assessment in educational systems, there are inadequate sources of information to apply for designing a comprehensive assessment. In a systematic literature review on evaluation of assessment conducted by Abdulghani and associates (2014)(8), it was concluded that most of the assessments are confined to the evaluation of a single

parameter such as reliability and validity of individual assessment instruments of MCQ rather than the overall assessment system (9). A well designed multiple choice questions (MCQs) are the preferred choice and mostly used test types since they are tenable to item analysis performed by computer software. There are numerous research reports on writing well-designed question items that attempt to measure the concept of interest. Through applying item analysis on the test items psychometric indices such as item difficulty, item discrimination, Cronbach alpha and best alternatives to the item response are determined. Despite the fact that designing MCQs are quite difficult to construct and fairly time-consuming even by the best trained instructors (10), applying MCQs in educational testing is a common practice. Considering the fact that this part of test evaluation is a prerequisite and necessary step, more comprehensive evaluation of a test is needed to determine what aspect of learning is assessed by constructing such test items. In cognitive domain of learning defined by Bloom (11), the main concern of assessment may include assigning of taxonomy to a question item, a practice that is not very often examined by the test makers. A well designed test is one that evaluates the cognitive levels in an education system (12-15).

Therefore, the aim of this study was to determine the taxonomy level of pediatrics exams and psychometric characteristics of MCQ test used in Kashan University of Medical Science in year 2014.

MATERIALS AND METHODS

In this descriptive cross-sectional research 150 multiple choice test items used to evaluate the pediatric residents at Kashan University of Medical Sciences in education year 2014 were item analyzed and the taxonomy of the test items were determined at three levels of recall-recognition, comprehension and application level (16-17) through an expert group of pediatric and educational psychology experts.

RESULTS

Item analysis was performed by computer software. Difficulty item (DI) was classified as very easy (DI = 0.85 to 1), easy (Di = 0.71 to 0.84), desirable level (DI =0.3 - .7) and very difficult (DI=0 to 0.29). From 150 question items, 42.7 percent were in recall and recognition (level 1), 31.3 percent were at comprehension level, and 26 percent were in application level. These results are presented in table 1. The result of chi-square test indicated that there were no significant differences among the proportion of taxonomy level of the test items (p=0.30).

Table 1: Frequency distribution of taxonomy level pediatric medical assistant exam

Taxonomy level	Frequency	percent	Cumulative Frequency
Recall-recognition	64	42.7	42.7
comprehension	47	31.3	74
Application	39	26	100
Total	150	100	-

The result of item analysis of pediatric medical assistant exam indicated that 35.3 percent of the test items were very easy (difficulty index between 0.85 to 1), 18 percent were easy (Di = 0.71 to 0.84), 30 percent were relatively at desirable level(di =0.3 to 0.7) and 13.3 percent were difficult(DI=0.16 to 0.29) and very difficult (DI=0 to 0.15). A significant differences among the proportion of difficulty level of the test items was found (p=0.000). These results are presented in table 2.

Table 2: Frequency distribution of question item difficulty levels

Difficulty Index	Frequency	Percent	Cumulative frequency
Very easy	53	35/3	35/3
Easy	27	18	53/3
Desirable	45	30	83/3
Hard	20	13/3	96/7
Very hard	5	3/3	100
Total	150	100.0	-

Further analysis was performed on the response alternatives to every question item. The result of analysis indicated that 32.7 percent of the question items had a single alternative chosen by the testers, 38 percent had two alternatives chosen, 22.7 had three alternatives chosen and only 6.7 percent of all on the alternatives were sufficiently

distracting for the test takers and were chosen as the correct alternative. There was a significant difference among the response alternatives being chosen as the correct alternative ($P=0.007$).

Table 3: Frequency distribution of response alternatives

Response Alternatives	Frequency	Percent	Cumulative frequency
1	49	32/7	32/7
2	57	38	70/7
3	34	22/7	93/3
4	10	6/7	100
Total	150	100	

DISCUSSION AND CONCLUSION

Multiple choice question tests are the preferred choice of exams and are commonly used to assess students learning in all level of education. Item analysis is a popular procedure to validate question items (3-5). This type of analysis is necessary for determining the appropriateness of every question items. However, this is not sufficient to conclude that a test is appropriate to assess students learning. A multiple choice question exam may include a large number of valid and reliable test items, but the items may address only one level of learning. For instance, a 100 item test that strictly targets recall and recognition not only lacks the property to assess higher level of learning but may also have poor diagnostic value to discriminate among the learners. These items are usually related to facts and concepts at first level and are easy to remember. Most students have the ability to memories them for testing purpose. In some instances, these items may be necessary to use since they are at the core of the learning level, but are not useful when higher level of learning is the target of testing. Therefore, level of learning a test measures is as important as the validity of the test as well as its reliability. Bloom (11) was one of the first authors who purposed the level of learning and introduced the taxonomy of educational domain. Ever since, taxonomy of educational domain has been the interest of researchers in education and educational psychology. The purpose of this research was to examine the item property of the multiple choice test as well as the level of learning purposed by Bloom (1956) in Pediatric medical assistants exam studying in Kashan University of Medical Sciences. The result of analysis showed that the question items included three levels of learning. From the total of 150 question items, 42.7 percent were in recall and recognition (level 1), 31.3 percent were at comprehension (level 2), and 26 percent were in application level (3). The result of chi-square test indicated that there were no significant differences among the proportion of taxonomy level of the test items ($p=0.20$). The result of this research was not in agreement with the findings of Abdolphani *et al.*(2015) (18) who reported that there was a significant differences among the learning levels of the faculty members when designed a multiple choice tests. The number of higher cognitive level of Bloom's taxonomy was significantly higher after the faculty members took part in a training course to design multiple choice question items. This finding reveals the significance of training of the faculty members to learn more about the level of learning as has been indicated by Bloom. However, the author did not report that what proportion of the question items were in higher cognitive level before the start of the research nor reports the proportions of different cognitive levels before or after the training course. The findings of the present research is in agreement with the findings of Blanco and associates (2015) and Légaré and associates (2015)(19-20)

The proportion of question items in the first level (recall and recognition) was considerably high but not statistically different. This is an indication that the test as a whole entity addressed easy items and addressed less high level of leaning.

Having test items that addresses low level of learning may encourage the learners to learn that level of learning(21-22) and conversely conducting test to measure higher level of cognitive functioning leads to adopting deeper method of training and learning (23).

In medical education higher levels of Bloom's taxonomy is needed. Levels that require demonstration of deeper cognitive processing such as critical thinking and evaluative judgments, but result of studies including the present research finding have shown that learning objectives in many training programs and curricula focus overwhelmingly on the lower levels of the taxonomy, knowledge and comprehension [19]. This is a serious problem for the medical education as well as other health related education since the graduates in the programs are expected to learn beyond solely memorized facts to be able to act for the prevention and treatment of their patients. In other hand, higher level of cognitive domain is necessary in medical education. It is very likely that the instructors and professors who are more or less specialists in different field of medicine need to participate in short term workshops or training courses to learn about the domain of learning in order to become competent in designing appropriate question item that addresses the assessment of higher cognitive level of learning. It needs to be noted that this study was limited to the

resident students in pediatric field. More research with other students at the general level of medical education and other departments will provide more comprehensive information about the quality of medical examinations in this university.

Acknowledgement

The researcher wishes to express his appreciation to the cooperation of Education Development Center of Kashan University of Medical Sciences in completion of this research.

Conflict of Interest: No Conflict.

Founding: This research was conducted by a grant from the research deputy of Kashan University of Medical Sciences.

REFERENCES

- [1] Amin Z, Khoo HE. In Basics in Medical Education. 2nd Edition, Singapore: Word Scientific Publishing; 2006. <http://dx.doi.org/10.1142/6109>
- [2] Van der Vleuten CPM. The assessment of Professional Competence: Developments, Research and Practical Implications. *Adv Health Sci Educ.* 1996;1(1): 41-67. <http://dx.doi.org/10.1007/bf00596229>
- [3] Sharif MR, Asadi MH, Sharif AR, Sayyah M, Examining the Multiple Choice Educational Examinations of College Students, *Biomedical & Pharmacology Journal* 6 (1), 23-27
- [4] Sayyah M, Vakili Z, Alavi NM, Bigdeli M, Soleymani A, Assarian M, An Item Analysis of Written Multiple-Choice Questions: Kashan University of Medical Sciences, *Nursing and Midwifery Studies*, 2012, 1 (2), 83-7
- [5] Sayyah M, Bigdeli M, Vakili Z, Assessing the Psychometric Indices of Written Multiple Choice Educational Examinations of College of Paramedics of Kashan University of Medical Sciences in Education Year 2008-09, *Biosciences Biotechnology Research Asia* 2011, 8 (Issue No.: 1), 125-129
- [6] Tarrant M, Ware J. Impact of item-writing flaws in multiple-choice questions on student achievement in high-stakes nursing assessments. *Med Educ.* 2008;42:198–206. <http://dx.doi.org/10.1111/j.1365-2923.2007.02957.x>
- [7] Tarrant M, Ware J, Mohammed AM. An assessment of functioning and non-functioning distractors in multiple choice questions: a descriptive analysis. *BMC Med Educ.* 2009;40(1):25-32. <http://dx.doi.org/10.1186/1472-6920-9-40>
- [8] Abdulghani HM, Ponnampereuma GG, Ahmad F, Amin Z. A comprehensive, multi-modal evaluation of the assessment system of an undergraduate research methodology course: Translating theory into practice. *Pak J Med Sci* 2014;30(2):227-232. doi: <http://dx.doi.org/10.12669/pjms.302.4372>
- [9] Schuwirth LW, van der Vleuten CP. A plea for new psychometric models in educational assessment. *Med Educ.* 2006;40(4):296-300. <http://dx.doi.org/10.1111/j.1365-2929.2006.02405.x>
- [10] Humaira Fayyaz Khan, Khalid Farooq Danish, Azra Saeed Awan, Masood Anwar, - Identification of technical item flaws leads to improvement of the quality of single best Multiple Choice Questions *Pak J Med Sci* 2013 Vol. 29 No. 3
- [11] Bloom B, Englehart M, Furst E, Hill W, Krathwohl D: Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. New York, Toronto: Longmans, Green; 1956.
- [12] Background Paper II: Entry-Level, Curricular Outcomes, Curricular Content and Educational Process; Commission to Implement Change in Pharmaceutical Education. *Am J Pharm Educ* 1993;57; 377-85.
- [13] Anderson L. W., Krathwohl D. R., Bloom B. S. A Taxonomy for Learning, Teaching, and Assessing a Revision of Bloom's Taxonomy of Educational Objectives. New York, NY: Longman; 2001
- [14] Zheng A. Y., Lawhorn J. K., Lumley T., Freeman S. Assessment: application of Bloom's Taxonomy debunks the "MCAT Myth." *Science.* 2008;319:414–415.
- [15] Alison Crowe, Clarissa Dirks, and Mary Pat Wenderoth Biology in Bloom: Implementing Bloom's Taxonomy to Enhance Student Learning in Biology
- [16] Kibble JD1, Johnson T. Are faculty predictions or item taxonomies useful for estimating the outcome of multiple-choice examinations? *Adv Physiol Educ.* 2011 Dec;35(4):396-401. doi: 10.1152/advan.00062.2011.
- [17] Thompson A.R , and O'Loughlin V.D., The Blooming Anatomy Tool (BAT): A discipline-specific rubric for utilizing Bloom's taxonomy in the design and evaluation of assessments in the anatomical sciences, *Anatomical Sciences Education* Volume 8, Issue 6, pages 493–501, November/December 2015 16 DEC 2014 DOI: 10.1002/ase.1507
- [18] Abdulghani, H.M. et al. Faculty development programs improve the quality of Multiple Choice Questions items' writing. *Sci. Rep.* 5, 9556; DOI:10.1038/srep09556 (2015).
- [19] Blanco MA, Capello CF, Dorsch JL, Perry GJ, Zanetti ML. A survey study of evidence-based medicine training in US and Canadian medical schools. *J Med Lib Assoc.* 2014 Jul;102(3):160–8. DOI: <http://dx.doi.org/10.3163/1536-5050.102.3.005>. [PMC free article] [PubMed]

- [20] Légaré F, Freitas A, Thompson-Leduc P, Borduas F, Luconi F, Boucher A, Witteman HO, Jacques A. The majority of accredited continuing professional development activities do not target clinical behavior change. *Acad Med.* 2015 Feb;90(2):197–202. DOI: <http://dx.doi.org/10.1097/ACM.0000000000000543>. [PubMed]
- [21] Ball A., Washburn S. Teaching students to think: practical applications of Bloom's taxonomy. *Agr. Educ. Mag.* 2001;74:16–17.
- [22] Bissell A. N., Lemons P. P. A new method for assessing critical thinking in the classroom. *BioScience.* 2006;56:66–72.
- [23] Gardiner L. F. *Redesigning Higher Education: Producing Dramatic Gains in Student Learning.* Washington, DC: George Washington University; 1994.